Communication, Control and Computing 563–570. Univ. Illinois.

SWENDSEN, R. H. and WANG, J. S. (1987). Nonuniversal critical dynamics in Monte Carlo simulations. *Phys. Rev. Lett.* **58** 86–88.

TAPLIN, R. and RAFTERY, A. E. (1994). Analysis of agricultural field trials in the presence of outliers and fertility jumps. *Biometrics*. **50** 764–781.

TIERNEY, L. (1994). Markov chains for exploring posterior distributions (with discussion). *Ann. Statist.* **22** 1701–1762.

TIERNEY, L. and KADANE, J. B. (1986). Accurate approximations for posterior moments and marginal densities. *J. Amer. Statist. Assoc.* **81** 82–86.

TIERNEY, L., KASS, R. E. and KADANE, J. B. (1989). Fully exponential Laplace approximations to expectations and variances of nonpositive functions. *J. Amer. Statist. Assoc.* **84** 710–716.

WEIR, I. S. and GREEN, P. J. (1994). Modelling data from single photon emission computed tomography. In *Statistics and Images* (K. V. Mardia, ed.) **2** 313–338. Carfax, Abingdon.

WHITTAKER, J. (1990). *Graphical Models in Applied Multivariate Statistics*. Wiley, New York.

WILKINSON, G. N. (1984). Nearest neighbour methodology for design and analysis of field experiments. In *Proceedings of the 12th International Biometrics Conference* 64–79. Biometric Society, Washington, DC.

WILKINSON, G. N., ECKERT, S. R., HANCOCK, T. W. and MAYO, O. (1983). Nearest neighbour (NN) analysis of field experiments (with discussion). *J. Roy. Statist. Soc. Ser. B* **45** 151–211.

WILLIAMS, D. (1982). Extra-binomial variation in logistic linear models. *J. Roy. Statist. Soc. Ser. C* **31** 144–148.

WILLIAMS, E. R. (1986). A neighbour model for field experiments. *Biometrika* **73** 279–287.

WRIGHT, W. A. (1989). A Markov random field approach to data fusion and colour segmentation. *Image and Vision Computing* **7** 144–150.

ZEGER, S. L. and KARIM, M. R. (1991). Generalized linear models with random effects: a Gibbs sampling approach. *J. Amer. Statist. Assoc.* **86** 79–86.

ZIMMERMAN, D. L. and HARVILLE, D. A. (1991). A random field approach to the analysis of field-plot experiments and other spatial experiments. *Biometrics* **47** 223–239.

Comment

Arnoldo Frigessi

In the beginning there was the Gibbs sampler and the Metropolis algorithm. We are now becoming more and more aware of the variety and power of MCMC methods. The article by Besag, Green, Higdon and Mengersen is a further step toward full control of the MCMC toolbox. I like the three applications, which show how to incorporate MCMC methods into inference and which also give rise to several methodological contributions. As the authors write, out of five main issues in MCMC, they concentrate primarily on the choice of the specific chain. The other four issues regard, in one way or another, the question of *convergence* of MCMC processes. I believe that choosing an MCMC algorithm and understanding its convergence are two steps that cannot be divided. Estimating rates of convergence (in some sense) before running the chain or stopping the iterations when the target is almost hit are needed operations if we would like to trust the inferential conclusions drawn on the basis of MCMC runs. This is especially true because convergence of MCMC processes is much harder to detect as compared to convergence of, say, Newton–Raphson.

Arnoldo Frigessi is Associate Professor, Dipartimento di Matematica, Terza Università di Roma, via C. Segre 2, 00146 Roma, Italy.

We can often read in applied papers that "100 iterations seem to be enough for approximate convergence," the number being sometimes supported by studies on simulated data (see, e.g., Frigessi and Stander, 1994). This is really too weak to rely on the statistical conclusions, and more can be done. If $X^{(t)}$ is the MCMC process with target distribution π on Ω, the *burn-in* can be estimated by computing a t^* such that

$$(1) \quad \forall\, t > t^*, \quad \|P(X^{(t)} = \cdot \mid x^{(1)}) - \pi(\cdot)\| \le \varepsilon,$$

for some fixed accuracy ε and for some chosen norm, say, total variation. Several techniques are available to bound the total variation error from above,

$$(2) \qquad \|P(X^{(t)} = \cdot \mid x^{(1)}) - \pi(\cdot)\| \le g(t),$$

where $g(t)$ is a nonincreasing function decaying to zero. Then an upper bound on t^* can be derived by inversion of g, probably a pessimistic estimate of the burn-in, but a "safe" choice. Tight bounds of the type (2) are hard to get and there are no precise general guidelines for the length of the burn-in. However a very *rough* reference value for t^* is available if π is a lattice-based Markov random field (MRF). In Section 1 of Frigessi, Martinelli and Stander (1993) we extend and adapt results originally developed in statistical mechanics and rather unknown to statisticians. Let π be a MRF on a